Vision Research 50 (2010) 107–117

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Graded recognition as a function of the number of target fixations

Michal Jacob *, Shaul Hochstein

Interdisciplinary Center for Neural Computation and Department of Neurobiology, Institute of Life Sciences, The Hebrew University, Jerusalem, Israel

ARTICLE INFO

Article history:
Received 3 May 2009
Received in revised form 28 October 2009

Keywords:
Eye movements
Recognition
Detection
Identity Search Task
Fixations
Visual awareness

ABSTRACT

Target recognition stages were studied by exposing observers to varying controlled numbers of *target* fixations. The target, present in half the displays, consisted of two identical cards (Identity Search Task; Jacob & Hochstein, 2009). Following more fixations, targets are better recognized, indicated by increased Hit-rate and detectability (according to Unequal Variance Signal Detection Theory), decreased Response Time and growing confidence, reflecting current stage in recognition process. Thus, gathering information over a *specific* scene region results from a growing number of fixations on that particular region. We conclude that several fixations on a scene location are necessary for achieving recognition.

1. Introduction

When searching for a complex target, what leads to detection? What are the stages and mechanisms in the process of target recognition? To study this issue, we examine the dynamics of detection and recognition as a function of the *number of fixations on the target*. We previously devised a novel Identity Search Task, where we found that target detection depends on the sequence of fixations (Jacob & Hochstein, 2009). We now use this task, again tracking eye movements, but halt the display after a certain (varied) number of fixations, specifically on the target, to catch different stages in the process of recognition.

There is a cognitive plan behind eye movements: fixation patterns or scan-paths are influenced by cognitive processes on the basis of task demands and the task-specific information available at different parts of the scene (Antes, 1974; Brandt, 1945; Buswell, 1935; Henderson & Hollingworth, 1999; Hochberg, 1970; Neisser, 1976; Ringach, Hawken, & Shapley, 1996; Yarbus, 1967).

In a recent example, eye movements were studied in the context of "change blindness". When viewing two alternating pictures of a scene, where a small or even a large difference has been introduced between them, observers are often "blind" to the change if it is not in the scene's focus of interest (Rensink, O'Regan & Clark, 1995, 1997; Simons & Levin, 1997, 1998; see Hochstein & Ahissar, 2002). It was found that even when the changing region has been fixated in both its states, detection rate is only 25% (Hollingworth, Williams, & Henderson, 2001). The possible need for several fixations per region was not examined.

Classical studies concluded that more informative scene regions receive more fixations (Antes, 1974; Buswell, 1935; Loftus & Mackworth, 1978; Mackworth & Morandi, 1967; Nodine, Carmody, & Kundel, 1978; Over, Hooge, Vlaskamp, & Erkelens, 2007; Yarbus, 1967) and that regions that receive more fixations are eventually identified (Nodine et al., 1978; they also found that Hits were preceded by examination-type, long-duration fixations, while Misses were preceded by survey-type, short-duration fixations; see also Over et al. (2007) who suggested that eye movements may follow a compulsory coarse-to-fine strategy). Are these extra fixations essential for recognizing targeted objects? Subjects in the experiments of Nodine et al. (1978) knew their target, (the word *Nina*), so that different regions of the display were more likely than others to be concealing the target. In our display, *a priori*, each region is as likely as any other to be the target. The only characteristic that renders one card a target is presence of another, identical card. This makes target detection quite complicated in our task, adding to task difficulty and trial duration – an advantage since examining the process is easier when it is slow. We do not compare local visual features in the scene, nor the semantics of the object (for example, consistent/inconsistent, Hollingworth et al., 2001). Rather, we examine how the number of fixations *on the target* influences the process of recognition.

We ask whether conscious target detection comes before or after extended target fixation. Do arbitrary multiple observations of the pair cards lead to detection, or, on the contrary, does

Abbreviations: CR, Correct Rejection; FA, False Alarm; ROC, Receiver Operating Characteristics; RT, Response Time; SDT, Signal Detection Theory.

* Corresponding author. Address: Department of Neurobiology, Institute of Life Sciences, Givat Ram, The Hebrew University, Jerusalem 91904, Israel.
E-mail address: michal.jacob@mail.huji.ac.il (M. Jacob).

0042-6989/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2009.10.019

detection of the pair lead to more fixations on it? In other words, does the larger number of fixations give rise to detection or does unconscious pre-recognition guide the eyes to more fixations?

There is evidence of a mismatch between fixation and detection, that is, that target fixation is not always accompanied by explicit detection (Barlasov-Ioffe & Hochstein, 2008; Hooge & Erkelens, 1996; Motter & Belky, 1998; Ruthishauser & Koch, 2007; Sheinberg & Logothetis, 2001). Recognition often requires a "double-take" saccade, i.e. one or more fixations away from the target (Ruthishauser & Koch, 2007), during which conscious recognition presumably occurs, leading the eyes back to the target.

We suggest that in the intervening period there is implicit perception (Mitroff, Simons, & Franconeri, 2002; Nodine et al., 1978; Rensink, 2004), which guides the eyes and saccadic planning. We further suggest that the following boosted target fixations bring this unconscious discovery to conscious awareness.

The *Identity Search Task*, which we first introduced in Jacob and Hochstein (2009), is a spatial recognition task, in which subjects are instructed to detect two identical cards (Fig. 1). The display contains computer screen "cards", each with a square array of scrambled black and white square units. The task is to detect two exactly identical cards, regarded as the target. The characteristics of the Identity Search Task that are important for our current research are that the identical card pairs do not pop out – rather, the recognition process requires several fixations on the target – and that displays are divided into distinct search and eye-fixation regions (the different cards) – allowing us to count fixations on the different regions. An enormous number of novel displays may be created, allowing us to repeat the task with a new search each time.

In our previous study we used 12-card displays, each with *two* pairs of identical cards, in order to compare between the eventually detected target and the undetected one – as we had done previously in a study of a more complex search task (Jacob & Hochstein, 2008). We found that the cards of the pair that was ultimately detected were observed more frequently than cards of the undetected pair. There were more fixations and longer fixations on the ultimately detected pair, and the average sequential distance between fixations on these card regions was smaller for the detected pairs. A bifurcation point was observed along the dynamics of search, in which the to be detected pair overpowered the undetected one.

In the main experiment of the current study there is only one identical pair, or none at all, and the task is not to actively detect the identical pair, but to state whether such a pair exists in the display at all. Eye Movements and fixations were recorded in real time to allow us to count the number of fixations on the pair cards, and to abort the display after a certain number. In this way, we controlled not the time of the display, but the more relevant parameter – the number of *target* fixations. We then analyze precision of detection response and degree of participant response certitude as a function of the number of target fixations achieved before the display is turned off. In this way we hope to measure the contribution of multiple fixations in the process of target detection.

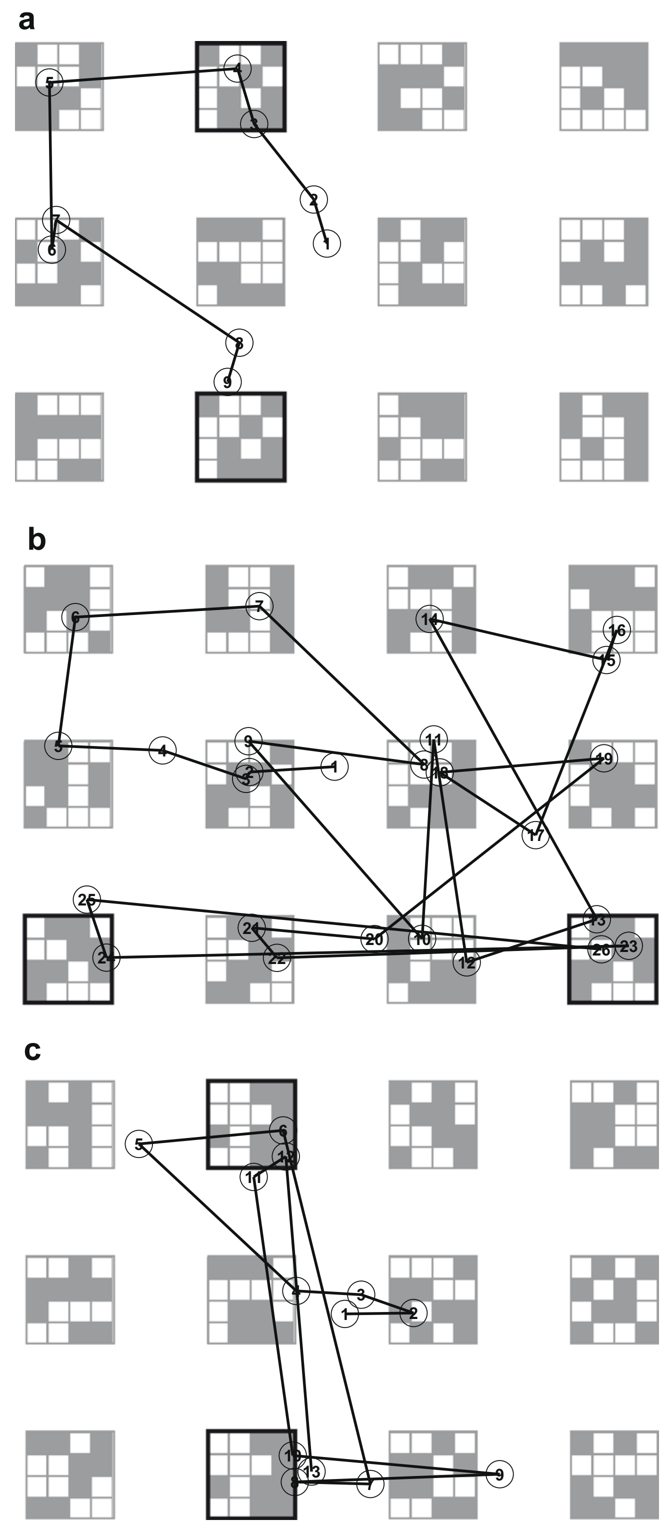

Fig. 1. The Identity Search Task. Twelve "cards" are displayed, each with a 4×4 scrambled array of black and white squares; subject task is to state whether they include an identical pair. Three examples are shown of displays with a target (shown here with a dark frame) with superimposed eye-movement records during search. For these three trials, the same subject had different numbers of target fixations before the display was turned off. The sequential number of each fixation is indicated in the circles. The Subject responded, "Yes" in all these examples, but reported different confidence levels. (a) Pre-determined number of target fixation was two, but incremented to three to obtain at least one fixation on each pair card. Confidence level: 'Do not know'. (b) Pre-determined and actual number of target fixations was five. Confidence level: 'Maybe'. (c) Pre-determined target fixations: seven; actual: six, i.e. an early response was given. Confidence level: 'Sure'.

2. Methods

2.1. The task

The experiment included three stages with the first two serving as training for the third, which included measurement of eye fixations and served as the central task of the experiment. Each stage included 100 trials and lasted about 30 min, with the third taking a bit more time to include calibration and drift correction of the eye movement monitor. For each Identity Search trial we presented a display of twelve "cards" and subjects were instructed to find two identical cards. Two identical cards did not appear in adjacent locations (i.e., one above the other, or side-by-side, but they could appear diagonally). The three stages were as follows:

1. Active detection of an identical pair in displays with exactly one identical pair, giving subjects practice and a sense of such displays. Subjects marked cards with mouse clicks. They could un-mark a card, as long as only one was chosen. Marking two different cards was considered a mistake, and a message 'Try again' appeared at the top of the screen. Response Time (RT) was measured from the appearance of the display until the click on the second correct card; subjects were informed of this timing procedure and that speed was important. At the end of a trial, a 'Continue' message appeared, and the next trial began when the subject mouse-clicked it. Then the entire display was replaced for the following trial.
2. Target/no-target training. In half of the trials there was an identical pair, and in half there was not. Fixations were not recorded and display times were randomized to approximately match a pre-determined number of fixations for each trial; (see below: Task Design). Subjects replied Yes|No according to presence|absence of a target, and they reported their confidence in this response, choosing from three possible levels (Don't know|Maybe|Sure).
3. Target/no-target + tracking eye movements (the core of the experiment). This stage was identical to the preceding one except that here we recorded eye movements and received real-time fixation data, allowing us to count target fixations (when a target was present) and to stop the trial after a pre-determined number of target fixations.

Nine subjects participated in the experiment. Two performed the third experimental stage twice, once with and once without reporting their confidence level (in either order). This was added to determine whether confidence level reporting caused a delay or change in initial Yes|No response.

The dominant eye of each subject was determined before the experiment, using the "hole-in-the-card" test (Durand & Gould, 1910; see review in Shneor & Hochstein, 2006). Fixations were analyzed according to the dominant eye.

2.2. Experimental routine for the central third stage

At the beginning of each trial, subjects were prompted with a message, "Press 'space' when ready". When the ready signal was given, we performed a drift correction using the center-of-screen point of the EyeLink built-in drift correction. Subjects were instructed to fixate carefully the dot before and while pressing the 'space' key, and were told that otherwise this may harm the experimental results. If after two presses of the 'space' bar, the dot did not disappear, meaning the drift was unacceptably large, which rarely occurred, we performed again the calibration, validation, and drift correction, and the experiment continued from that point. After successful disappearance of the drift correction point, the display was shown.

For trials with a target, the display was turned off after a pre-determined number of fixations on it, and the mask was presented. As per prior instructions, subjects responded by pressing the 'g' or 'h' key, for 'No' or 'Yes', respectively. (Red and green stickers were attached to these keys.) Subjects were instructed that response correctness was primary and speed secondary. After responding 'g' or 'h', subjects were prompted with the message, "Confidence: 1 – Don't know; 2 – Maybe; 3 – Sure", and were given as much time as needed for replying. They were allowed an early Yes|No response before the display disappeared and they were then prompted for a confidence level reply. They pressed the 'space' bar to begin the next trial.

2.3. Payment procedure

Subjects received, a basic payment of NIS 50 for the whole session (~$13), plus, for the third stage, a bonus of NIS 0.50 for each correct response (Hit or Correct Rejection) above chance level and for correct above incorrect early responses (disregarding net negative bonuses).

Subjects were informed of this payment procedure, were given examples, and were explained that the optimal tactic would be to try to give a correct response, and only then, to attempt to give it quickly. The objective of this payment procedure was to encourage as early a response as possible, once subjects knew the answer, given that they could not know when the display would disappear. In this way we obtained information regarding the number of fixations required for explicit recognition.

2.4. Analysis of number of target fixations

We combined successive fixations on the same card region (See Discussion in Jacob & Hochstein, 2009), whether they were target or non-target card regions, if: (1) The distance between them was less than 2° (67 pixels) *and* (2) One of them lasted <130 ms *or* the two together lasted <330 ms. More than two fixations could be combined if each pair obeyed these conditions. The durations of the combined fixations were summed. (Nevertheless, eye-movement records in Fig. 1 reflect uncombined fixations.)

2.5. Design of the main task

In the target/no-target eye-movement task, an identical pair target appeared pseudo-randomly in half of the trials. The number of target fixations was randomized in advance, from the range of 2–7 and 10 fixations (this range was determined by the distribution of detected pair cards fixations in our earlier experiment; Jacob & Hochstein, 2009). When this number of fixations was reached, the trial was aborted and a mask was presented. In displays with a target pair, we required also that there be at least one fixation on each of the pair cards. Otherwise, the trial was continued until the subject made a fixation on the other card.

Subjects were not informed that when we tracked eye movements we also aborted trials following a certain number of fixations; they were told that termination time was random.

Four considerations guided determination of the distribution of number of fixations:

1. *Combinatorial calculations*: For each number of fixations n, there are 2^n possible divisions of the fixations between the two cards (counting also opposite scenarios), but two of these have all fixations on the same card. To compensate for these, we multiply the pre-planned number of trials with the desired number of target fixations by $2^n/(2^n - 2)$. For example, the multiplication factors for number of trials with 2, 3, 4 and 5 target fixations were 2, 4/3, 8/7 and 16/15, respectively. As the number of target fixations grew, the multiplication factor became negligible. However, see restriction in clause 2, explaining why the formula is fully applied only to the 2-target fixations trials.
2. As mentioned, if the pre-determined number of fixations was reached with all of them on one card, we waited for at least one fixation on the second card of the pair. Thus, the number of target fixations was higher than intended (See example in Fig. 1a). Therefore, such a trial will be added to one of the groups of a bigger number of target fixations. This allows pre-

determination of a smaller number of trials for groups of target fixations greater than 3.

3. If an early response was given, then the number of target fixations was lower than intended (See example in Fig. 1c).
4. We combined successive fixations on the same card, as elaborated above. This led to a reduction in actual number of target fixations.

All of the above factors interact to influence the pre-determined number of trials in each group. Taking all of these factors into consideration led to the approximation used and shown in Table 1 (pre-determined) together with the average actual resulting number – to which we relate in all the following analyses. Due to the low number of trials with 10 target fixations, the results were analyzed only for the target fixation range of 2–7.

For trials without a target, we obviously could not count target fixations; therefore, each trial was assigned a duration (t), which was randomized according to the matching number of fixations (n) and using a linear regression of the means ($t = [1.2 \cdot n + 1.67]$ s) calculated from a previous experiment. This was jittered randomly in the range of $t \pm 1$ s to avoid biasing subjects. False Alarm rates were related to this equivalent target fixation number.

When a display with a target resulted in no fixations on the target at all, or fixations on just one of the target cards, it follows that the subject gave an early response. This could happen in two ways: (1) The subject did not respond correctly, i.e. responded 'No', when actually a target was present. (2) The subject considered two other cards as a pair; (otherwise, the subject must have perceived the cards with peripheral vision, which is not very likely).

2.6. Implementation

The experiment was implemented using the GUI of Matlab 7.0.4. Cards were represented as clickable buttons, uniformly distributed over three rows and four columns. Each card occupied 85 × 85 pixels (~2.7° × 2.7° of visual field) with a 92 (horizontal) and 86 (vertical) pixel space between cards. The borders of the *regions* used for analysis were taken at half this distance, including for peripheral cards, (a radius of <2°; Anstis, 1974; Riggs, 1965). The mouse pointer was moved off screen.

2.7. Equipment

Dominant eye fixations were recorded with an SR Research Ltd. (Ontario, Canada) EyeLink I eye-tracker. Subjects sat constrained by a chin-rest, 80 cm from a Samsung SyncMaster 19 in. CRT monitor, with 4:3 format and screen resolution of 800 × 600 pixels so that the foveal field of 2° occupies 2.8 cm or 67 monitor pixels. The monitor was surrounded by a black screen.

We used the SR-supplied (binocular) 9-point calibration and validation grid, repeating as necessary (effective radial resolution was 0.6 deg.) Drift correction was performed before each trial. The EyeLink was controlled by Matlab Psychophysics and Eyelink Toolboxes (Brainard, 1997; Cornelissen, Peters, & Palmer, 2002). Fixation and saccade analyses were performed with the EyeLink program. 'End-fixation' events were retrieved in real time to count the number of target fixations.

Table 2
Mean distribution of responses.

Response	Target present	Target absent	Total (subject control)
Yes	**23 Hits**	6 FAs	29
No	27 Misses	**44 CRs**	71
Total (pre-determined)	50	50	100 Trials/subject

3. Results

3.1. Performance

For 100-trial tests, the average total number of correct answers (Hits plus Correct Rejections, CR) was 67 (range 61–74), where 50 was chance level. On average, subjects responded early in 15 of 100 trials (range 4–46; after an average of five fixations), of which 11 (44–100%) were correct. Average bonus payment was NIS 12 (range NIS 6–19).

Table 2 shows the across-subject mean distribution of response types – Hits, Misses, Correct Rejection (CR) and False Alarms (FA). For 50 target and 50 no-target displays per subject, there were only 29 'Yes' answers, on average, and 71 'No' replies, indicating conservative strategies overall. This is to be expected since we stopped the trials – and turned off the display – often at quite early stages of search. Subjects may well be saying "No, I did not detect the target" rather than "No, I am sure there is no target." This is to our advantage, since we are interested in the process of detecting the target. The fact that subjects were all quite conservative means that they answered "No Target" by default and needed to be convinced (even if implicitly) that there was a target. We follow this process of "being convinced" as it proceeds with display time and number of target fixations.

Still, some subjects were more conservative than others. There were two distinct groups of responses: five more conservative subjects gave 14–21 'Yes' answers, including 0–2 FAs, and four less conservative subjects gave 35–46 'Yes' answers, including 7–14 FAs. Interestingly the two strategies led to the same average number of correct responses (18 Hit + 49 CR vs. 30 Hit + 37 CR).

3.2. Number of target fixations

The core of this research is to investigate the influence of the number of target fixations on target recognition. We show the dependence on number of target fixations of correct responses (Hits and CRs; Fig. 2a), mean RT (Fig. 2b) and confidence level (Fig. 2c), as well as the effect on resulting Signal Detection Theory ROC curves (Fig. 5).

Performance (Hit, Miss, FA and CR) is plotted in Fig. 2a as a function of actual number of target fixations (or their equivalent for

Table 1
Distribution of average number of trials for each number of target fixations, when present, or equivalent duration when target absent.

Present: # fixations (Absent: time ±1 s)	**1**	**2** (4.1)	**3** (5.3)	**4** (6.5)	**5** (7.7)	**6** (8.9)	**7** (10.1)	**8**	**9**	**10** (13.7)	Sum
Target Absent											
Planned and actual #	0	**8**	7	7	7	7	7	0	0	7	50
Target present											
Pre-determined #	0	**14**	**5**	**6**	**6**	**6**	7	0	0	**6**	50
Actual #	0.4	7.7	**9.1**	**8.9**	**7.9**	**6**	**5.8**	1	0.8	2.3	50
Actual # (with combined fixations)	0.8	**9.6**	**10.3**	**10.1**	7.3	**5.9**	**3.3**	1.1	0.9	0.7	50

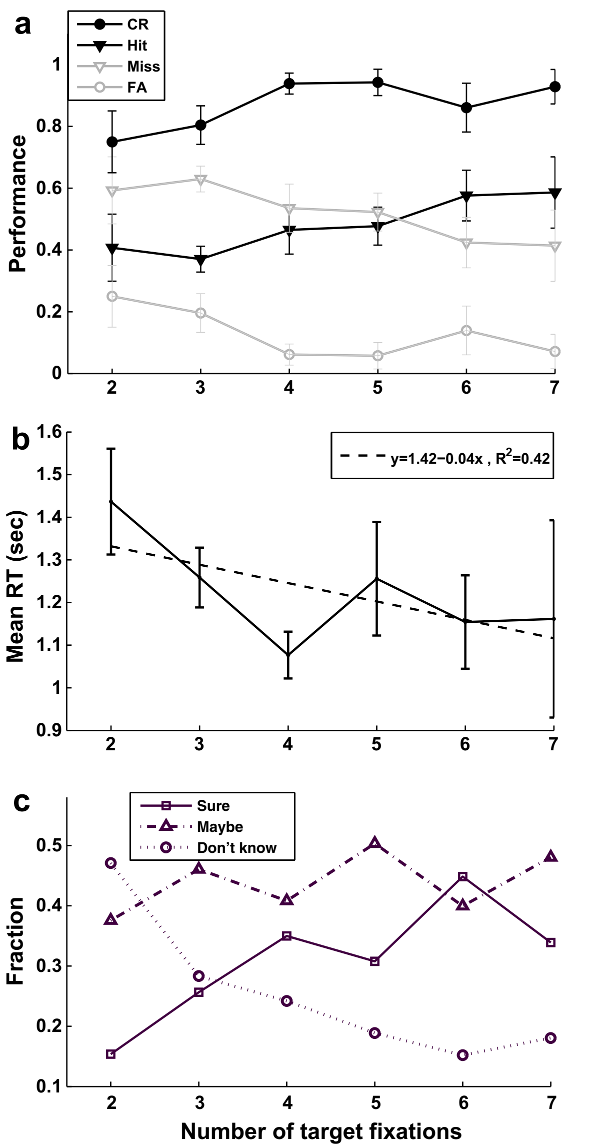

Fig. 2. Impact of target fixations. (a) Performance (rates of Hit, Miss, FA and CR) vs. number of target fixations; mean over all nine subjects (900 trials). Error bars are between-subject *SE*. (b) Mean RT (between subjects) after disappearance of display vs. number of target fixations, not including trials with an early response. Error bars represent between-subject *SE*. Nine subjects, 765 trials (excluding 135 early responses). (c) Fraction of each confidence level as a function of number of target fixations (for 450 displays with a target). Note sharp increase in 'Sure' rate at the expense of decreasing 'Don't knows' (and approximately constant 'Maybe' responses) with increasing target fixations.

no-target displays; see Section 2). Note the trend for increasing Hit-rate and decreasing FA-rate with number of target fixations (known as the Mirror Effect; DeCarlo, 2007; Glanzer & Adams, 1990; Glanzer & Bowles, 1976; Wixted, 1992). When going from 2–3 fixations to 6–7 fixations, there are more Hits (39% → 57%) and fewer FAs (23% → 10%), as shown in Fig. 3, making for a large increase in the *adjusted* Hit-rate (i.e. correcting for guesses by subtracting the FA-rate from the Hit-rate, with the result going from 16% to 47%). We performed an ANOVA on the adjusted Hit-rate with main factors of number of target fixations (combining data for 2–3, 4–5, and 6–7 target fixations) and subject (as a random factor) and found significance for the dependence on number of target fixations ($F = 8.04$, $p = 0.004$), but not for subjects ($F = 0.98$, $p = 0.486$). Post hoc paired *t*-tests over subjects, showed an increase in Hit-rate for 6–7 target fixations compared to that for 2–3 target fixations ($p < 0.05$) and a decrease in FA-rate for the same

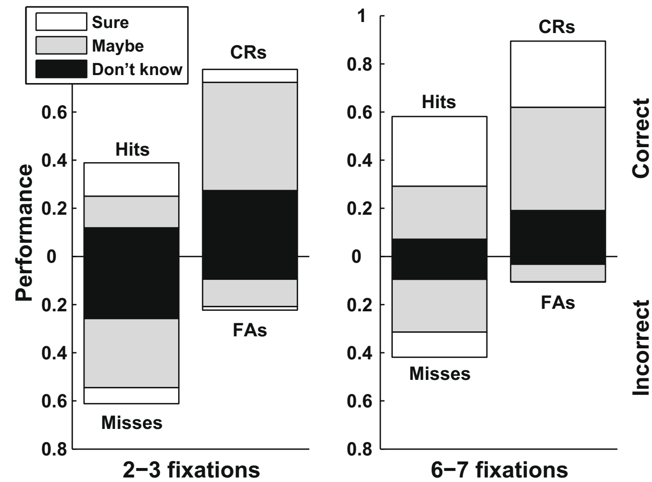

Fig. 3. Histogram representing division of performance (Hits, Misses, CRs and FAs) into confidence levels. Correct responses are shown above 0 and incorrect responses below; responses relating to target present and target absent displays are shown in the left and right columns, respectively. Performance in displays with a target (Hits, Misses) sum to 1, as does performance without a target (CRs, FAs). Shown is the progress in performance and surety from 2–3 target fixations (left graph) to 6–7 target fixations (right graph).

number of fixations ($p < 0.05$). We conclude that the number of target fixations does significantly influence detection.

Mean Response Time (RT) from display disappearance to Yes|No response was 1.3 ± 0.2 s (between subject mean ± *SD*) for both target present and target absent trials, but, of course, excluding trials with an early response when the display was not extinguished by the experimenter. Post-display-disappearance RT falls with the number of fixations, as shown in Fig. 2b, probably reflecting the increased information that subjects have concerning target presence or absence.

It would seem that a fixation, which lasts ~200 ms 'saves' only 40 ms in Response Time, which might seem not 'worthwhile', but, actually, each fixation raises the amount of information available, as seen in the performance and confidence results in Fig. 2a and c (confidence results are discussed in the following paragraph); RT serves only as an indicator of information, not as the goal in itself.

We now look at confidence level as a function of target fixations, demonstrated in Fig. 2c. Confidence level responses turned out not to be 100% reliable in the sense that subjects were often "sure" but wrong. Nevertheless, there is still a consistent increase in surety with number of fixations. Fig. 2c shows the fraction of each of the three confidence levels as a function of the number of target fixations for target present trials. 'Maybe' responses are constant at ~45% while 'Sure' responses rise from 15% to ~40% at the expense of decreasing 'Don't know' responses. We further use a *weighted average surety index*, weighting the confidence response according to: Don't know – 0, Maybe – 0.5, Sure – 1. When going from 2–3 target fixations to 6–7 target fixations, the Surety index (see Table 3) increased (Hits surety 0.52 → 0.73; CR surety 0.36 → 0.56; Miss surety 0.33 → 0.56; FA surety 0.32 → 0.36; paired *t*-test over subjects, $p < 0.001$. There were almost no 'Sure' responses following very few target fixations (in the CRs, Misses, and FAs: 6–8%, for 2–3 target fixations; not shown). As the number of target fixations increases, 'Sure' responses rise for Hits, and with further increase in the number of fixations, 'Sure' responses rise also for CRs and Misses (not shown). We conclude that with more target fixations, subjects become more confident of their answers.

We analyzed the data with Signal Detection Theory (SDT; Green & Swets, 1966; for a recent review, see Wixted, 2007). The SDT distributions in our case are "signal" – a display with a target (the

Table 3
Distribution of confidence level across subjects.

Subject	**Don't know**	**Maybe**	**Sure**	'Yes' responses	Hits	Correct	**Weighted average – surety**
Y.S.	**50**	46	4	14	13	63	0.27
O.H.	**52**	38	10	46	30	64	0.29
T.G.	**55**	28	17	19	17	67	0.31
A.A.	38	31	31	39	25	61	0.47
J.H.	12	**78**	10	35	32	69	0.49
N.L.	8	**67**	25	20	20	70	0.59
E.T.	20	40	40	38	31	74	0.60
I.G.	0	**70**	30	18	16	64	0.65
D.K.	2	35	**63**	21	20	69	0.81

Table is sorted by weighted average surety index. Interleaving number of 'Yes' responses suggest that there is no correlation between conservatism and surety.

right Gaussian, with mean μ_2 and standard deviation σ_2) – and "noise" – a display without a target (the left Gaussian, with mean μ_1 and standard deviation σ_1).

Fig. 4 presents ROC and zROC curves for each number of target fixations, averaged over trials. The curves were constructed from points plotted for each of the five criteria separating the different confidence level responses. That is, each point along a ROC or zROC curve represents the cumulative Hit- and FA-rates for each of the six different responses, ordered from 1-"No-Sure" and 2-"No-Maybe" to 5-"Yes-Maybe" and 6-"Yes-Sure", beginning with the most confidently recognized pairs (i.e. Hit-rate = P[6|target]; FA-rate = P[6|no target]) and repeatedly recalculating the rates by including the next most confidently recognized pairs (Egan, 1958; See review by Yonelinas & Parks, 2007). This procedure results in shifts of the criterion to the left of the distributions (or upward-and-to-the-right in ROC space) when moving from high to low confidence in 'Yes' responses and from low to high confidence in 'No' responses. The sixth point is constrained to be 1, as it complements all type of responses, and is not shown. Curves for each number of target fixations were plotted in Fig. 4a according to mean and standard deviation of the distribution inferred from the regression in the *z*-space.

The same data are plotted in Fig. 4b in *z*-space, the inverse of the standard cumulative normal distribution (assuming a mean of 0 and standard deviation of 1), and linearly regressed for each number of target fixations. When a zROC is linear, the *slope* equals σ_1/σ_2, (where σ_1 and σ_2 are the standard deviations of the "noise", i.e. target absent, and "signal", i.e. target present distributions, respectively), so a slope <1 indicates greater variance of the 'signal' distribution and that the data do not match an Equal Variance Signal Detection model. The zROC *intercept* equals $(\mu_2-\mu_1)/\sigma_2$ and, when divided by the slope, yields $(\mu_2-\mu_1)/\sigma_1$ – the distance between the means of the signal and noise distributions in units of the noise distribution standard deviation (which is assumed fixed; Martini & Maljkovic, 2009) – an Unequal Variance model analog of the Equal Variance Signal Detection model detectability d'. Slopes and intercepts of the zROCs are plotted in Fig. 4c and d. Note that both change mainly from 2–3 to 4–5 fixations, accompanied by an increase in detectability, as seen in the curves of Fig. 4a.

The likelihood ratio is the probability that an observation came from the signal distribution divided by the probability that it came from the noise distribution. In this sense, optimal placement of the criterion is at the point where the two probabilities are equal, i.e., where the likelihood ratio = 1. As the criterion shifts left, the likelihood ratio increases, and vice versa. The dashed diagonal in Fig. 4a denotes the optimal criterion separating 'Yes' and 'No' responses.

We show the log likelihood for all five separating criteria, and for each number of target fixations, in Fig. 4e. The inset shows likelihood ratio for the criterion separating between 'Yes' and 'No' responses (as presented in the middle column of Fig. 4e), vs. number of target fixations.

Moving from 2–3 to 4–5 and then to 6–7 target fixations, the likelihood ratio of the criterion separating 'Yes' and 'No' responses increases (Fig. 4e-inset), indicating a higher Hit-rate to CR-rate ratio, therefore implying a shift to the left of the criterion, resulting in more 'Yes' responses, that is, less conservatism (though the ratio is still smaller than 1). Fig. 4e shows that the log likelihood ratios represented by all the five separating criteria become less sparse with an increase in number of target fixations. That is, in the cross-section of six target fixations (black star), for instance, the points are less spread apart than in the cross-section of two target fixations (gray circles). This indicates that with more target fixations subjects respond 'Sure' not just in the extremities, therefore shortening the intervals of the 'Don't know' and 'Maybe' responses (as all the criteria become closer to the middle of the distribution space). This means that subjects become more confident in their responses with more target fixations.

3.3. Number of fixations as an indicator of the stage in the process of recognition

To learn about the effect of the number of fixations on the process of recognition, we use confidence level as an indicator of the stage in this process. 'Don't know' may reflect a situation of 'searching in the dark', where the response is just a guess; when responding 'Maybe', subjects already have a vague idea (conscious or unconscious), but they are still not sure – this is a situation of implicit recognition without confidence; 'Sure' is equivalent to the stage of full perception. We show how the number of fixations dictates the stage in the process of recognition in Fig. 5.

We plot, for each level of confidence, the normalized average number of target fixations that led to a Hit. That is, for each confidence level (c), we average the number of fixations (f) that led to a Hit, weighted by the relative number of Hits at that confidence level, or

$$\sum_f \left(\frac{Hit_{f,c}}{Hit_f} \cdot f\right) \Big/ \sum_f \frac{Hit_{f,c}}{Hit_f}.$$

This is the same as calculating the center of gravity of the surface below the plots in Fig. 2c.

There is an increase in number of fixations from level to level of confidence. Results are shown in Fig. 5 for seven subjects (two were excluded because of low Hit-rates despite declaring 'Sure'). Increase is steepest for subjects whose confidence reports were more reliable; (reliability is demonstrated in Fig. 7a). After an average of three fixations they are still in the stage of "searching in the dark". After an average of five fixations, they already have full perception of the target. Somewhere in between they are in the implicit recognition stage. Subjects' report of a 'Maybe' confidence level seems to reflect a sense of the presence of a target, without explicit knowledge. We infer that progress along the process of recognition requires added fixations.

3.4. Confidence level

Having shown that the number of target fixations affects performance levels and speed (Fig. 2) and confidence level (Figs. 2 and 5), we now investigate the relationship between confidence level itself and the rate of correct responses. Fig. 6 shows a performance histogram (Hits, Misses, CRs and FAs) divided into confidence levels. A number of characteristics are evident: there are many fewer 'Yes' than 'No' responses for our conservative observers. Examination of target (Hits + Misses) and no-target (CRs + FAs) displays reveals that subjects responded 'Don't know' in about 26% (13 of 50)

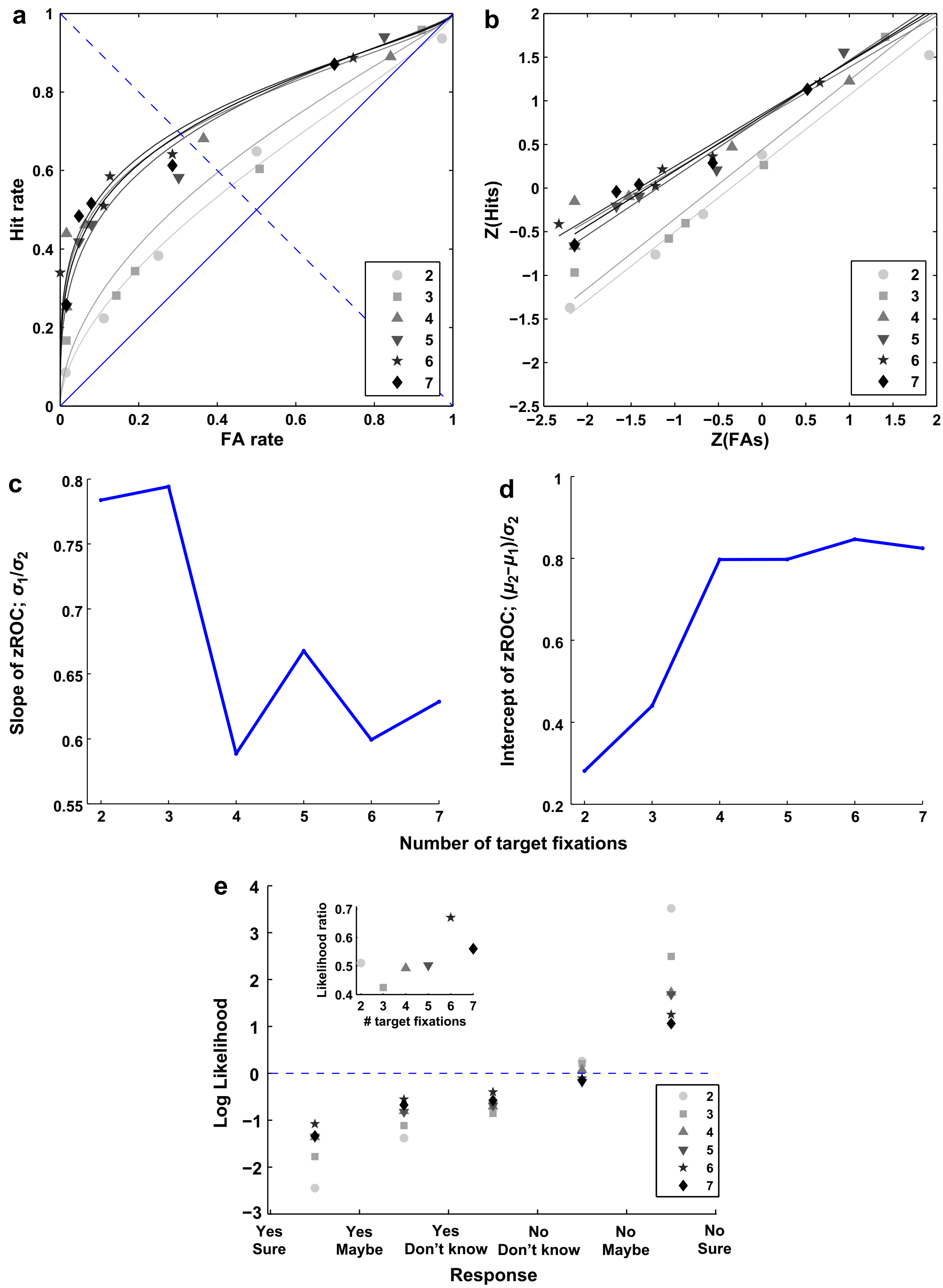

Fig. 4. (a) Receiver Operating Characteristics (ROC) curves for each number of target fixations (indicated by gray scale and marker); mean over trials. ROCs were constructed from the five criteria separating responses, such that each point along a ROC curve represents cumulative Hit- and FA-rates of the six different responses (i.e. Yes|No × Don't know|Maybe|Sure). (b) zROCs; data from a plotted in *z*-space. (c) Slope of zROCs as inferred from a linear regression, indicating σ_1/σ_2, vs. number of target fixations. (d) Intercept of zROCS, indicating $(\mu_2-\mu_1)/\sigma_2$, vs. number of target fixations. (e) Log likelihood for all five separating criteria, and for each number of target fixations; *inset:* likelihood ratio for the criterion separating between 'Yes' and 'No' responses vs. number of target fixations.

of trials in either case. 'Sure' responses were more prevalent for correct replies (Hits and CRs), i.e., observers were rarely sure when they were wrong. The main difference in confidence distribution between correct-response trials without vs. with a target (CRs vs.

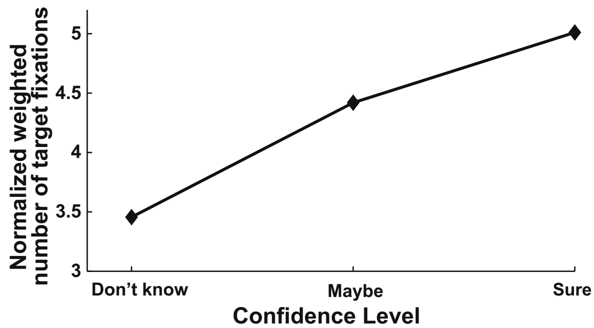

Fig. 5. Normalized weighted average number of target fixations for trials with Hits for each confidence level. There is an increase in number of fixations needed to reach each level of confidence. Data for seven subjects.

Hits) was a shift from the most common response, "maybe" to the "sure" response for displays with a target. This might be expected since one can be sure of having seen a target, but only lengthy systematic scanning can make one sure of target absence.

We examined the direct effect of confidence level on performance (Hit, Miss, FA and CR), as shown in Fig. 7a. Hit-rate increases with confidence level, and FA-rate slightly decreases. RT also decreases as confidence level increases, as demonstrated in Fig. 7b. Early responses were excluded; data are for seven subjects because two did not give any 'Don't know' or any non-early 'Sure' responses. The clear RT decrease confirms the accuracy of subjects' confidence level self-report.

To determine whether the need to declare confidence level influenced the Yes|No responses themselves or their RTs, we tested two subjects twice, once with and once without reporting their confidence level (counter-balancing the order of the two runs). We found no change in their pattern of responses, consistent with their own strategy, (number of Yes|No responses; number of Hits, Misses, FA and CRs), but there was a major speeding of their RTs

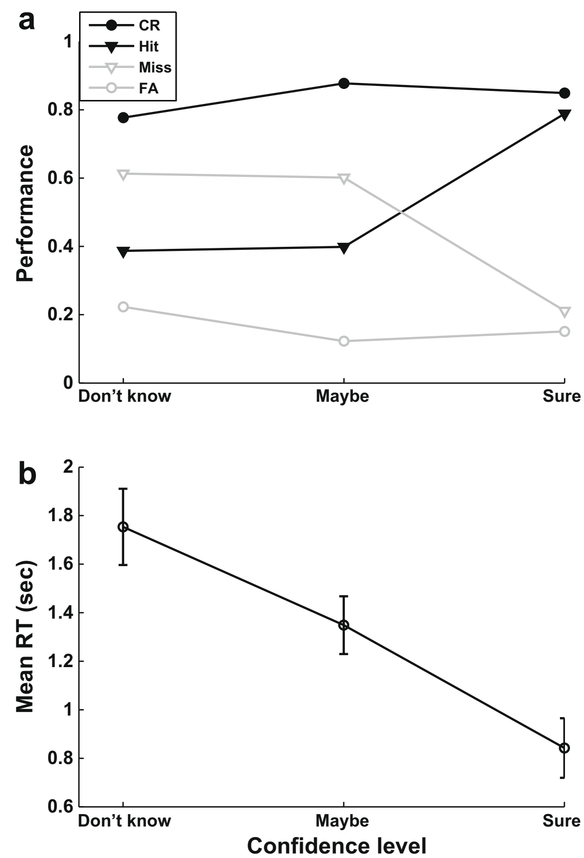

Fig. 7. (a) Performance (Hit-, Miss-, FA- and CR-rates) vs. confidence level. Note rising correct responses with confidence level. (b) Mean (between subjects) after-display-disappearance RT vs. confidence level. Note decline in RT with confidence level. Compare Figs. 2a and b, respectively.

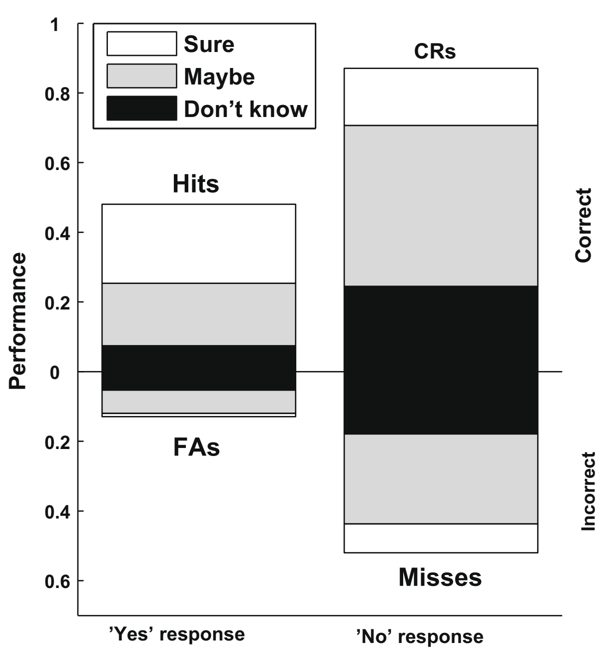

Fig. 6. Performance histogram (Hits, Misses, CRs and FAs) divided into confidence levels. Correct responses are shown above 0; incorrect responses, below; 'Yes' responses are shown in the left column and 'No' responses on the right. Note that performance in displays with a target (Hits, Misses) sum to 1, as do performance without a target (CRs, FAs). 'Sure' responses are found mainly for correct replies (Hits, CRs), and they are more prevalent (with a concomitant decrease in 'Maybe' response) for target present Hits than target absent CRs.

from 1.5 s with confidence level reporting to 1.0 s without. This was true for both subjects and irrespective of order, suggesting it was not a learning effect. (Again, there was no difference between target- and no-target displays). Perhaps the need to declare one's confidence level caused hesitation and delay in giving the Yes|No response, but it did not affect the contents of that response. There was also no influence on number of early responses.

Individual subject confidence level distributions are shown in Table 3. The order of the rows is according to weighted average Surety (with weights: Don't know: 0; Maybe: 0.5; Sure: 1). It would seem that different strategies for declaring confidence level were adopted by different subjects. Note that confidence level is not directly linked to criterion, i.e. to degree of subject conservatism: one might expect that a less conservative subject would also be more confident in his or her response, but it turned out not to be so. For example, a subject who gave only 21 'Yes' responses is the one who gave the highest number of 'Sure' responses. Less conservative subjects, by definition, indeed gave more 'Yes' answers, but apparently they were aware of that they were 'gambling', and declared they were not sure. The confidence level declaration allowed subjects to respond 'Yes' even if they were not sure, or even did not know at all.

3.5. Impact of non-target fixations

Is there a relationship between detection and the fraction of fixations that are on the target? Fig. 8 shows the Hit-rate as a function of the total number of fixations for a fixed number of target fixations. That is, as the total number increases, the fraction on the tar-

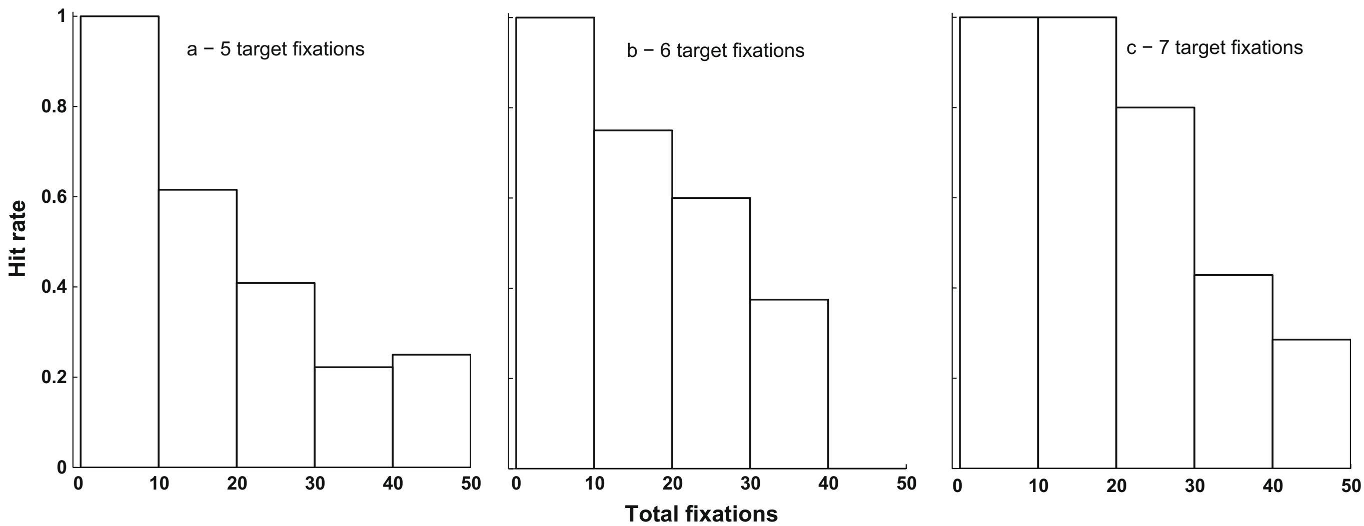

Fig. 8. Hit-rate as a function of different total number of fixations on the display for the cases with 5–7 target fixations. Note declining Hit-rate with increased *total* number of fixations, for fixed number of target fixations.

get decreases. Clearly, the Hit-rate decreases with decreasing target fixation fraction – even though the number of target fixations is fixed. We conclude that target fixations need to be closer together and/or not disturbed by distracting fixations on non-target cards. The main effect of many display fixations is not a contribution to familiarity; it is distraction from the target.

4. Discussion

Our aim was to understand better different stages in the process of detecting and recognizing a target and to determine the effect of the number of target fixations on the transition from stage to stage. To this end, we exposed subjects to controlled numbers of target fixations – by stopping the display at different points – and inferred subject recognition stage from their performance and confidence levels.

The target, if present, comprised two identical cards in the Identity Search Task. There were two responses: a timed Yes|No response relating to target presence or absence and a report of the subject's level of confidence in that response, as Don't know| Maybe|Sure.

Evidence for the influence of the number of target fixations comes from a number of our experimental findings, as follows:

1. Improved performance. Hit-rate increases as a function of the number of target fixations (Fig. 2a) and there is an increase in detectability and a decrease in the Yes–No criterion towards optimality (Fig. 4). This is consistent with earlier research, which revealed that when two targets are present, the one that is detected is that with more fixations on it (Jacob & Hochstein, 2009). Subjects were not very good in recognizing the target after very few fixations on it.
2. Faster response with increasing target fixations, implying that an increase in number of target fixations leads to progress in the recognition process (Fig. 2b).
3. More confident responses. As the number of target fixations increased, subjects became more confident of their responses (Fig. 2c).

Clearly, the number of target fixations necessary for recognition depends also on the type of target itself. In our experiment, because the target consisted of two separate cards, at least one (or even two) fixation(s) on each card were necessary for recognition, hence the poor results for total of only two and three fixations.

Unequal Variance SDT analysis showed that the zROC intercepts increase and slopes decrease from 2–3 to 4–7 target fixations (Fig. 4d,c) leading to increased detectability (Fig. 4a). The increased asymmetry stems from the decrease in slope, which reflects the ratio σ_1/σ_2, indicating a broadening signal distribution and increasing difference in noise and signal distributions (i.e. a greater tendency towards an *Unequal* Variance Signal Detection model.

As for the criterion, it is by definition placed where the likelihood ratio is <1 for conservative subjects. Moving from 2–3 to 4–5 and then to 6–7 target fixations, the likelihood ratio of the criterion separating between 'Yes' and 'No' responses increases (Fig. 4e-inset), implying a shift of the criterion to the left along the recognition axis, resulting in more 'Yes' responses, that is, less conservatism. In general, the five separating criteria become closer together with increased number of target fixations (Section 3.2 and Fig. 4e), implying more confident responses.

Interestingly, detectability increases mainly from lowest to mid-level number of target fixations; there is initially an increase in d' and then stabilization. Detectability does not improve from 4 to 7 fixations, but subjects are still improving in their criterion.

Moving from 2–3 to 4–7 target fixations, the distributions separate, as seen in the increased ratio of intercept to slope, $(\mu_2-\mu_1)/\sigma_1$ and in accord with the mirror effect (DeCarlo, 2007; Glanzer & Adams, 1990; Glanzer & Bowles, 1976; Wixted, 1992), and the signal distribution becomes relatively broader. The less-than-unity slope supports an Unequal Variance Signal Detection model (see Jang, Wixted, & Huber, 2009) and the *change* in variance and detectability suggests a Mixture Signal Detection model (DeCarlo, 2002, 2007; Jang et al., 2009), in which the target distribution is a mixture of two Gaussians, with different means – lower for unattended or partially attended items (in our case, 2–3 fixations), and higher for attended items ($\geqslant$4 fixations).

There is a clear difference between the 2–3 target fixation ROC curves and the 4–7 target fixations ROC curves. We therefore conclude that a major change has occurred as a result of more target fixations.

Misses may occur despite many target fixations, when target fixations are interspersed with many non-target fixations, which may be both distracting, (due to the larger sequential distance between target fixations), and give the sense that if the target has not been spotted by now, it might not be present. Thus, there is an inverse influence of total number of fixations over the entire display on detection (Section 3.5 and Fig. 8). This rules out the option that the described effects are due just to search time, and not to number of target fixations. Performance improves with more target fixations, but for a fixed number of target fixations, performance declines with more fixations on the entire display, that is, with increased search time.

In Jacob and Hochstein (2009) we proposed a 3-stage model of the perceptual recognition process during visual search: *Stage 1*: An initial "search in the dark", consisting of fixations in pseudorandom order (i.e., in our task – not dependent on card pattern or at least not *a priori* on its belonging to a matched pair); *Stage 2*: Implicit (unconscious) detection of the target, guiding further eye movements to the target location, i.e. boosting fixations on it; *Stage 3*: Explicit detection with conscious knowledge of target presence and its location, dependant on crucial fixations on the target.

To learn about the effect of the number of target fixations on the process of recognition, we use confidence level as an indicator of the stage in the process. The 'Don't know' level is regarded as equivalent to the 'search in the dark', the 'Maybe' level to the stage of implicit recognition, and the 'Sure' level to the stage of full perception (explicit recognition). We find that there is an increase in the mean number of target fixations corresponding to the transition from one level of confidence to the next (Fig. 5).

We turn to the question of what comes first, more fixations or recognition. If only a vast number of target fixations can lead to recognition, we would not expect the rise in Hit-rate that was found after very few fixations. So we conclude that even a few fixations lead to some recognition. Yet, this partial recognition, reflected in the Hit-rate, which is above chance level but not very high, is then followed by more fixations, until reaching complete recognition (early responses or "sure" confidence). Uniting these two observations, we suggest that a few fixations lead to an implicit recognition state, which in turn leads to more fixations guided to the relevant sensed location of target cards, leading eventually to full explicit recognition.

Further support for stages along the process of recognition comes from the presence of two steps in the SDT model. There is an initial increase in detectability and asymmetry of distributions; a transition occurring with the rise from 2–3 to 4–7 target fixations. Following this transition, there is another shift in the separating criterion between 'Yes' and 'No' responses, when rising from 4–5 to 6–7 target fixations. Additional evidence for a gradual process comes also from: (1) The different performance, i.e. Hit-rate, accompanying the different levels of reported confidence level (Fig. 7a); (2) Change in the separating criteria, expressed by the likelihood ratio, between the different confident responses (Fig. 4e); (3) Decrease in RT when moving from one confidence level to another (Fig. 7b).

We suggest that recognition is a gradual process, as evident by gradually increasing degrees of both accuracy and confidence (see Mickes, Wais, & Wixted, 2009), as opposed to a categorical process, in which accuracy is high only with complete confidence and for all other confidence ratings accuracy is no better than chance (forming a step function relationship between accuracy and confidence).

A remark regarding the Aha! experience in detection (Ahissar & Hochstein, 1997; Bowden & Jung-Beeman, 2007; Maier, 1931; Rubin, Nakayama, & Shapley, 1997; Smith & Kounios, 1996; Sternberg & Davidson, 1995) – it might seem that if the process of recognition is gradual, as we suggest, then there is no momentary experience of discovery. This is not true, because the process of recognition includes also the stage of implicit recognition, in which there is no awareness of the discovery (See Ahissar & Hochstein, 1997, 2004; Hochstein & Ahissar, 2002). Therefore, even though a gradual process is taking place, transfer from unconscious to conscious recognition can emerge momentarily.

It has recently been claimed that there is no complete representation of the visual scene built up and remembered across visual fixations. Instead, memory depends on return fixations to recall what is in previously visited sites (O'Regan, 1992). Additional claims of this sort were based on phenomena such as change blindness – not noticing that details are changed between fixations (e.g. Rensink, 2000; see also Hollingworth et al., 2001) or slow (serial) visual search being independent of fixed element position (Horowitz & Wolfe, 1998). Hollingworth et al. (2001) indeed tested if change blindness would prevail following single fixations on the changing region, once before and once following the change. They find that in only 25% of the cases did this suffice for change detection. However, the conclusion that there is no inter-saccade memory would not follow if information were gathered gradually over a number of fixations to a scene region – as we now find. Thus, it would be of interest to repeat Hollingworth et al.'s (2001) experiment, but testing change detection following multiple fixations on the changing region. We would predict that detection will rise with number of fixations to the region before and after the change – and with the temporal proximity of such fixations.

We suggest that it may be more economical for the visual system to scan the visual scene and gather partial information from each sampled region, instead of expending all its resources on one location to gather full information from one site at a time. Attention can be spread uniformly at first, gathering minimal information in minimal time, and only then repeatedly to already observed regions, (perhaps emphasizing more important locations), to accumulate information gradually. The advantage of this scanning method comes from creation of parallel partial representations for all locations, rather than full knowledge about one location, on account of none about others.

We conclude that graded recognition derives from graded information gathered fixation after fixation on the same scene region.

5. Conclusions

We exposed subjects to controlled stages along the recognition process by varying the number of target fixations. We found that with more target fixations there is an increase in Hit-rate and detectability (d'), and in relative signal distribution variance; a decrease in RT, perhaps indicating greater availability of information; a decrease in conservatism, indicated by the shifts to the left of the Yes–No separation criterion; and an increase in confidence levels indicated by the observers and closer separating criteria for all response levels. Taken together, these results imply improved recognition with more target fixations. That is, an increase in number of target fixations leads to progress along the recognition process, and thus to more accurate, faster, and more confident reactions.

We found an inverse influence of total number of fixations over the entire display on detection. When target fixations are interspersed with many non-target fixations, those intervening fixations are distracting, perhaps due to the larger sequential distance between target fixations.

All these findings support the conclusion that gathering of information over a *specific* region of the scene results from a growing number of fixations on that particular region. This leads to the conclusion that several fixations on a scene location are necessary for achieving recognition.

Acknowledgments

This study was supported by grants from the Israel Science Foundation (ISF) and the US–Israel Binational Science Foundation (BSF). We thank the Charles E. Smith Family and Prof. Joel Elkes Laboratory for Collaborative Research in Psychobiology for use of its eye tracking equipment. We thank Reut Avinun for her contribution to the experiment. We thank SR Research Ltd. for their advice and technical support. We thank anonymous reviewers for their suggestions.

References

Ahissar, M., & Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. *Nature, 387*, 401–406.

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences, 8*, 457–464.

Anstis, S. M. (1974). Letter: A chart demonstrating variations in acuity with retinal position. *Vision Research, 14*, 589–592.

Antes, J. R. (1974). The time course of picture viewing. *Journal of Experimental Psychology, 103*, 62–70.

Barlasov-Ioffe, A., & Hochstein, S. (2008). Perceiving illusory contours: Figure detection and shape discrimination. *Journal of Vision, 8*(11), 1–15. 14.

Bowden, E. M., & Jung-Beeman, M. (2007). Methods for investigating the neural components of insight. *Methods, 42*, 87–99.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.

Brandt, H. F. (1945). *The psychology of seeing*. New York: Philosophical Library, Inc.

Buswell, G. T. (1935). *How people look at pictures*. Chicago: University Chicago Press.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The eyelink toolbox: Eye tracking with MATLAB and the psychophysics toolbox. *Behavior Research Methods, Instruments and Computers, 34*, 613–617. <http://cornelis.med.rug.nl/pub/EyelinkToolbox>.

DeCarlo, L. T. (2002). Signal detection theory with finite mixture distributions: Theoretical developments with applications to recognition memory. *Psychological Review, 109*, 710–721.

DeCarlo, L. T. (2007). The mirror effect and mixture signal detection theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 13–33.

Durand, A. C., & Gould, G. M. (1910). A method of determining ocular dominance. *Journal of the American Medical Association, 55*, 369–370.

Egan, J. P. (1958). *Recognition memory and the operating characteristics*. (United States Air Force Operational Applications Laboratory Technical Note Nos. 58, 51, 32).

Glanzer, M., & Adams, J. K. (1990). The mirror effect in recognition memory: Theory and data. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 5–16.

Glanzer, M., & Bowles, N. (1976). Analysis of the word frequency effect in recognition memory. *Journal of Experimental Psychology: Human Learning and Memory, 2*, 21–31.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology, 50*, 243–271.

Hochberg, J. (1970). Attention, organization, and consciousness. In D. I. Mostofsky (Ed.), *Attention: Contemporary theory and analysis* (pp. 99–124). New York: Appleton-Century-Crofts.

Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron, 36*, 791–804.

Hollingworth, A., Williams, C. C., & Henderson, J. M. (2001). To see and remember: Visually specific information is retained in memory from previously attended objects in natural scenes. *Psychonomic Bulletin and Review, 8*, 761–768.

Hooge, I. T., & Erkelens, C. J. (1996). Control of fixation duration in a simple search task. *Perception and Psychophysics, 58*, 969–976.

Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature, 394*, 575–577.

Jacob, M., & Hochstein, S. (2008). Set Recognition as a window to perceptual and cognitive processes. *Perception and Psychophysics, 70*, 1165–1184.

Jacob, M., & Hochstein, S. (2009). Comparing eye movements to detected vs. undetected target stimuli in an identity search task. *Journal of Vision, 9*(5), 1–16. 20.

Jang, Y., Wixted, J. T., & Huber, D. E. (2009). Testing signal-detection models of yes/no and two-alternative forced-choice recognition memory. *Journal of Experimental Psychology: General, 138*, 291–306.

Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance, 4*, 565–572.

Mackworth, N. H., & Morandi, A. J. (1967). The gaze selects informative details within pictures. *Perception and Psychophysics, 2*, 547–552.

Maier, N. R. F. (1931). Reasoning in humans. II. The solution of a problem and its appearance in consciousness. *Journal of Comparative Psychology, 12*, 181–194.

Martini, P., & Maljkovic, V. (2009). Short-term memory for pictures seen once or twice. *Vision Research, 49*, 1657–1667.

Mickes, L., Wais, P. E., & Wixted, J. T. (2009). Recollection is a continuous process: Implications for dual-process theories of recognition memory. *Psychological Science, 20*, 509–515.

Mitroff, S. R., Simons, D. J., & Franconeri, S. L. (2002). The siren song of implicit change detection. *Journal of Experimental Psychology: Human Perception and Performance, 28*, 798–815.

Motter, B. C., & Belky, E. J. (1998). The guidance of eye movements during active visual search. *Vision Research, 38*, 1805–1815.

Neisser, U. (1976). *Cognition and reality*. San Francisco: Freeman.

Nodine, C. F., Carmody, D. P., & Kundel, H. L. (1978). Searching for Nina. In J. W. Senders, D. F. Fisher, & R. A. Monty (Eds.), *Eye movements and the higher psychological functions* (pp. 241–257). Hillsdale, NJ: Erlbaum.

O'Regan, J. K. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology, 46*, 461–488.

Over, E. A., Hooge, I. T., Vlaskamp, B. N., & Erkelens, C. J. (2007). Coarse-to-fine eye movement strategy in visual search. *Vision Research, 47*, 2272–2280.

Rensink, R. A. (2000). Seeing, sensing, and scrutinizing. *Vision Research, 40*, 1469–1487.

Rensink, R. A. (2004). Visual sensing without seeing. *Psychological Science, 15*, 27–32.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1995). Image flicker is as good as saccades in making large scene changes invisible. *Perception, 24*(Suppl.), 26–28.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8*, 368–373.

Riggs, L. A. (1965). Visual acuity. In C. H. Graham (Ed.), *Vision and visual perception* (pp. 321–349). New York: Wiley.

Ringach, D. L., Hawken, M. J., & Shapley, R. (1996). Binocular eye movements caused by the perception of three-dimensional structure from motion. *Vision Research, 36*, 1479–1492.

Rubin, N., Nakayama, K., & Shapley, R. (1997). Abrupt learning and retinal size specificity in illusory-contour perception. *Current Biology, 7*(7), 461–467.

Ruthishauser, U., & Koch, C. (2007). Probabilistic modeling of eye movement data during conjunction search via feature-based attention. *Journal of Vision, 7*(6), 1–20. 5.

Sheinberg, D. L., & Logothetis, N. K. (2001). Noticing familiar objects in real world scenes: The role of temporal cortical neurons in natural vision. *The Journal of Neuroscience, 21*, 1340–1350.

Shneor, E., & Hochstein, S. (2006). Eye dominance effects in feature search. *Vision Research, 46*, 4258–4269.

Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences, 7*, 261–267.

Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people in real-world interaction. *Psychonomic Bulletin and Review, 5*, 644–649.

Smith, R. W., & Kounios, J. (1996). Sudden insight: All-or-none processing revealed by speed–accuracy decomposition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1443–1462.

Sternberg, R. J., & Davidson, J. E. (1995). *The nature of insight*. Cambridge, MA: Bradford Books/MIT Press.

Wixted, J. T. (1992). Subjective memorability and the mirror effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 681–690.

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review, 114*, 152–176.

Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum.

Yonelinas, A. P., & Parks, C. M. (2007). Receiver operating characteristics (ROCs) in recognition memory: A review. *Psychological Bulletin, 133*, 800–832.